

Genature version 5.1.3
Copyright (c) 1992 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:47:13 / Search time 18 seconds

(without alignments)
1474 061 Million cell updates/sec

Title: US-09-855-294b-1

Perfect score: 1438

Sequence: 1 MTSFPIALQLQINQIFSN... QIFVPPQKFSLYAVEMVFP 276

Scoring table: PROSUM62

Gapop 10.0, Gapext 0.5

Searched: 292224 seqs, 9613402 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 01

Maximum Match 100%

Listing first 45 summaries

Database: PIP 71.4

1: pirl:*

2: pirl2:*

3: pirl3:*

4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query %	Match	Length	DP	ID	Description
1	814	56.6	318	2	148563		hypothetical prote
2	773	53.8	322	2	128399		probable amidhydr
3	712	49.5	272	2	969109		N-carbamoyl p-amin
4	694.5	48.3	291	2	551459		hypothetical prote
5	450.5	31.3	440	2	143198		nitrilase/phit pro
6	418	29.1	271	2	AC2056		hypothetical prote
7	406.5	28.3	307	2	56607		hypothetical prote
8	398	27.7	275	2	582325		conserved hypothet
9	397.5	27.6	270	2	581199		nitrilase NME0441
10	397	27.5	264	2	136488		probable hydrolase
11	396	27.5	283	2	267362		carbon-nitrogen hy
12	395	27.5	284	2	AD3275		beta-uridylproton
13	392.5	27.3	272	2	576224		hypothetical prote
14	388.5	27.0	270	2	581834		conserved hypothet
15	383.5	26.7	316	2	DA5088		nitrilase i like p
16	381	26.5	286	2	AE2398		amidhydrolase (im
17	381	26.5	293	2	719215		hypothetical prote
18	368.5	25.6	282	2	E83086		conserved hypothet
19	358	24.9	283	2	A00446		fixable carbon ni
20	355	24.7	276	2	140601		putative nitrilase
21	354.5	24.7	295	2	EA2556		beta-alanine synth
22	353.5	24.6	259	2	E69863		conserved hypothet
23	351.5	24.4	262	2	C75051		hydrolase related
24	347.5	24.2	291	2	E95106		carbon nitrogen hy
25	343.5	23.9	262	2	C71109		hypothetical prote
26	342.5	23.8	231	2	C97204		beta-uridylproton
27	339	23.6	264	2	AC1470		conserved hypothet
28	335.5	23.3	277	2	719264		conserved hypothet
29	330	22.9	269	2	AC1110		conserved hypothet

30	324	22.5	340	2	C70743	hypothetical prote
31	316.5	22.2	347	2	E65063	fixable hydrolase
32	316	22.0	271	2	E87015	probable hydrolase
33	307	21.3	280	2	T24905	probable hydrolase
34	306.5	21.3	294	2	AP0115	probable carbon-ni
35	306	21.3	290	2	E81329	fixable hydrolase
36	303	21.1	271	2	G83780	hypothetical prote
37	300	20.9	280	2	E65064	hypothetical prote
38	300	20.9	292	2	G83608	probable hydrolase
39	297	20.7	294	2	G71940	hypothetical prote
40	293	20.4	292	2	E65558	conserved hypothet
41	292.5	20.3	292	2	E71501	hypothetical prote
42	290.5	20.2	292	2	E64514	beta-alanine synth
43	283.5	19.7	261	2	S20703	hypothetical prote
44	282.5	19.6	261	2	H00004	hypothetical prote
45	279.5	19.4	271	2	H93195	conserved hypothet

ALIGNMENTS

RESULT 1

T48563

hypothetical protein F14F18.210 Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20 Apr 2003 #sequence_revision 10 Apr 2000 #text_change 01 Feb 2002

C:Accession: T48563

P:Revan, M., Hiltner, H., Braun, M., Helzer, E., Brandt, A., Eusterhoefft, A.; Bandoft,

submitted to the Protein Sequence Database, April 2000

A:Reference number: 234490

A:Accession: T48563

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-318 <BEV>

A:Cross-references: EMBL:AL163812

A:Experimental source: Cultivar Columbia; EAC clone F14F18

C:Genetics:

A:Map position: 5

A:Introns: 25/3, 64/3, 120/3, 127/1, 207/1, 231/2, 236/3, 256/3, 261/2

A>Note: F14F18.210

C:Superfamily: hypothetical protein YLR351c

Query Match: 56.6%, Size 814, DB 2, Length 318;

Best Local Similarity: 56.6%; Pred No. 6 le-63;

Matches 102, Conservative 95, Mismatches 72, Indels 16, Gaps 2;

QY 1 MTSFPIALQLQISSPKSCNVTACGIPDPFAACGAGTIVLPECPHFYGCAPYFVAVY 60

DE 23 LTVENGLCLLSLTLPFFNINSHAPFAIEAAKELVLLPFIWISFYANSEFFVAAE 82

QY 61 I-----PGESTOKLSEVAKECSLYLGSSIPPEPACPIYNTQAVPGDGTILAPYPIHL 115

DE 83 IDATGTRASPSTAMLEVSVPKPTITGSIPEPVGSLVNTQVPGDGLKAVHPVHL 142

QY 116 FQIVGKPTIFESKTSLSPESTSTTPY-----TSWELITVMPFARACI 164

DE 143 FEFIFSEPIFMSPLTAGETITVIGMLSEPHIFGSEIGTICVIFQELKAM 202

QY 165 YAPFGQGLLVYFAPNITTPFAPHWELIAPSPFAVENAVVATAPAPFVAVVAVHSTV 224

DE 202 YAFJAHLLVYFAPNITTPFAPHWELIAPSPFAVENAVVATAPAPFVAVVAVHSTV 202

QY 225 VHWSEVLAAGTEATVSTITVFIAFTSLEITVFETFTLY 269

DE 263 VGFEGEVLAATHEEAILIAEIVSILEQFTGLFNPQFEGDLY 307

RESULT 2

T38399

probable amidhydrolase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 01 Dec 1993 #sequence_revision 01 Dec 1993 #text_change 01 Feb 2000

C:Accession: T38399

GenBank version 5.1.1.3
Copyright (c) 1991-2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on January 31, 2003, 07:44:12 ; Search time 11 seconds
(without alignments)
1040 679 Million cell updates/sec

Title: US-09-855-294b-1

Perfect score: 1438
Sequence: 1 MTSFPLALQLAIISSIKSN

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Search: 112892 seqs, 41476328 residues

Total number of hits satisfying the search criteria: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	EB	ID	Description
1	773	53.8	322	1	YAUB_GCHPO	Q10166 schizosacch
2	694.5	48.3	291	1	YLAE_YEAST	P49964 sacccharomyc
3	406.5	28.3	307	1	YLMK_YEAST	P47016 sacccharomyc
4	392.5	27.3	272	1	YECI_STH1?	P55175 synochocyst
5	324	22.5	340	1	Y480_MYCTU	Q11146 mycobacteri
6	283.5	19.7	261	1	Y4G5_STAAU	P55177 staphylococ
7	263.5	19.3	252	1	YBEM_ECOLI	P19874 escherichia
8	257.5	17.9	234	1	YAGE_STALU	P55178 staphylococ
9	256.5	17.8	262	1	YBEM_ECOLI	P58054 escherichia
10	254	17.7	339	1	NPLI_ARATH	P12962 arabidopsis
11	247.5	17.2	513	1	YHGX_BACSU	P54608 bacillus su
12	244	17.0	346	1	NPLI_ARATH	P12961 arabidopsis
13	240.5	16.7	346	1	NPLI_ARATH	P46010 arabidopsis
14	239	16.6	295	1	YFOQ_ESEEL	P55176 pseudomonas
15	206	14.3	349	1	NPLI_ARATH	P42965 nicotiana
16	203.5	14.0	365	1	NPLI_ARATH	Q03317 rhodospirillum
17	202.5	14.1	355	1	NPLI_ARATH	P46011 arabidopsis
18	175	12.2	346	1	ALAM_PSEAE	P11436 pseudomonas
19	173	12.0	256	1	YAFV_ECOLI	Q47679 escherichia
20	162.5	11.3	345	1	ALAM_PSEAE	Q01360 rhodococcus
21	155	10.9	393	1	NPLI_PSEAE	P02068 rhodococcus
22	153.5	10.7	520	1	YNNI_HUMAN	Q05498 homo sapien
23	152	10.6	393	1	RUPI_RAT	Q03248 rattus norv
24	152	10.6	501	1	YNNI_HUMAN	Q05498 homo sapien
25	151	10.5	500	1	YNNI_MOUSE	Q09425 mus musculu
26	148.5	10.3	512	1	YNNI_MOUSE	Q09425 mus musculu
27	149	10.3	576	1	NAE2_THEME	Q09425 mus musculu
28	142.5	9.9	514	1	YNNI_GAFA	Q09425 mus musculu
29	140	9.6	343	1	NPLI_PSEAE	Q05498 homo sapien
30	145	9.4	344	1	RUPI_HUMAN	Q05498 homo sapien
31	134	9.3	457	1	NTAI_YEAST	P40354 sacccharomyc
32	133	9.2	368	1	YNNI_GLOSO	P12964 arabidopsis
33	129.5	9.0	439	1	LNT_AQUAE	Q67006 aquifex aeo

RESULT 1
YAUB_SCHPO STANDARD; PRT: 322 AA.
AC 210.76,
DT 01-OCT-1996 (Pel 34, Created)
DT 01-OCT-1996 (Pel 34, Last sequence update)
DT 15 JUN 2002 (Pel 41, Last annotation update)
DE Hypothetical protein C26A3.11 in chromosome I.
GN SPAC26A3.11.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetes; Schizosaccharomycetes;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
PP SEQUENCE FROM N.A.
PC STPAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam P., Rajandream M.A., Lyne M., Lyne P., Stewart A.,
Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris P., Hidalgo J., Hodgson G.,
Holroyd S., Holsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Murgall M., Murphy L., Niblett D., Odell C.,
Olliver P., O'Neill P., Pearson D., Seal M.A., Sefton D., Smith E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellern J., Simmonds M., Squares P., Squares S., Stevens K.,
Taylor V., Taylor P.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Wolchert G., Aert P., Pothier J., Grynolprez B.,
Welfjens I., Vanstreels E., Pieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs V., Fritz C., Holder E., Moesl D., Hilbert H.,
Borzym K., Langer I., Beck A., Lebrach H., Weinhard P., Pohl T.M.,
Eger P., Giermann W., Wedler H., Warburton F., Dunnally B.,
Geffeau A., Cadieu E., Dreano S., Gloux S., Jelaire V., Mottier S.,
Galibert F., Aves S.J., Xiang Z., Hunt G., Moore K., Hurst S.M.,
Cass M., Foster M., Tallard M., Tallard M.A., Glick A., Thibault J.,
Raga P.F., Choudhury R., Dimerer J., Sanchez M., Al-Pay F., Benard J.,
Domenger A., Revuelta J.L., Morero S., Armstrong J., Forburg S.B.,
Cerrito L., Lowe R., McCombie W.R., Paulsen O., Patschkin J.,
Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
The genome sequence of Schizosaccharomyces pombe.
PL Nature 415 971-980(2002)
CC SIMILARITY: BLASTX TO THE UPF0012 FAMILY.

ALIGNMENTS

Q0427 homo sapien
C67991 aquifex aeo
C0411 h's filia
Q01638 rhodobacter
P43551 homo sapien
P23633 differella
P74055 synochocyst
P40447 sag-lai-myc
Q91139 neisseria m
P74992 synochocyst
P71311 mycobacteri
P20960 alcaligenes

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isit.sib.ch/infocenter> or send an email to license@isib-sib.ch.

EMBL: D69443; CAA0334.1;
InterPro: IPR00010; N150/ThyA

[illegible]

Db 62 ESTAVKELAEJLVAVAGMFTADTVQGEKTSVNNNTVLISG--AGLHQGYNKIHTY 119
QY 117 EIDVPGKIPQCKYKTSQSGSSTPNPYFVGLGICYMPPFAELAGIYVQKGCQLLVYF 176
Db 120 P--AFYPSSTVPGDHEIUVFVDIDIFFGVATYDIRFPEQPKDLARNGAQIIVVP 174
QY 177 GAFNLTTGPA--HWELLQSPAVDQVQVYVATASPA-----RDDKASYVAMGHSTVVNP 227
Db 175 TSMQ--DGPQFLQVAVLPAPALLSTWTIVAGCGAPLPBELRDERKGTGIGHSVMVNP 232
QY 228 WSEVLAFATHEAIVTSIDPKYLAIEPQCLPV 260
Db 233 HGEVIASAGYEPDMITADIDVSLGLAKIREALPV 265

RESULT 2
US-09-774-639-188
; Sequence 188, Application US/09774639
; Publication No. 19990203035A;
; GENERAL INFORMATION
; APPLICANT: Posen et al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: P0303P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; PRIOR FILING DATE: 2001-07-09
; PRICE APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRICE FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver 2.0
; SEQ ID NO 188
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-188

Query Match 23.1%; Score 332.5; DB 9; Length 153;
Best Local Similarity 45.2%; Pred No 2 5e-26;
Matches 66; Conservative 25; Mismatches 48; Indels 7; Gaps 3;
QY 128 ESYTSLQDQSF STPTTPYFVGHGICYDMPFAELAGIYVQKGCQLLVYVQAEHLTTGPA 186
Db 3 ESNSTMPGFSLESFVSTPAGVIGLAVCYDMPFPHLSIALAQAGAEILTYPSAFGSITGPA 62
QY 197 HWELLQSKAVLRGVVYVATASPA---ELDKASYVAMGHSTVVVFWGSEVLAKVTEEAIVY 243
Db 63 HWEVLRLARA-ETQCYVVAQAQCGRHHEKRASY---GHSWVVDPMGTIVVARGCEGPGGL 119
QY 244 SDIDLYFLAEIETQVPTVQFQPSJLY 269
Db 120 ARIDINYLQFLRHLPLVPQHRPPDLY 145

RESULT 3
US-09-972-186A-4
; Sequence 4, Application US/09972186A
; Patent No. US20020137153A1
; GENERAL INFORMATION:
; APPLICANT: PAMER, SANDRA W.
; APPLICANT: HUISMAN, GJALT
; APPLICANT: MILLIS, JIM
; APPLICANT: SHELDON, ROGER
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: COX, ANTHONY
; APPLICANT: DAVIS, S. CHRISTOPHER
; TITLE OF INVENTION: ENANTIOSELECTIVE PRODUCTION OF AMINO CARBOXYLIC ACIDS
; FILE REFERENCE: 02-108110US
; CURRENT APPLICATION NUMBER: US/09/972,186A
; PRIOR FILING DATE: 2001-10-04
; PRICE APPLICATION NUMBER: 60/238,563
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide sequence
US-09-972-186A-4

Query Match 14.0%; Score 201.5; DB 10; Length 366;
Best Local Similarity 24.8%; Pred No. 1.5e-12;
Matches 80; Conservative 51; Mismatches 124; Indels 67; Gaps 15;
QY 3 SPRLALIQIQ-ISSIKSDNVTRACSFIREAATOCAKIVSLPECF ... -NSPYG- 50
Db 7 TPKVAAVQAPQVWFDAKTVCKIVSIIAEARNNGCELVAPEVFIPGYPVHIWVDSPLAG 66
QY 51 AKYFEYAEK...IFESTQKLESEVAECETILIGTSIFEDACGLYNTAVPQFETTL 106
Db 67 MAKFAVRYHENSUTMDSPHVQKLLDAARHSAIVAVG-ISEKLGSSLMVQLIIDAGGQL 125
QY 107 LAKYRKHLPLDIDVPGKILTHQESKILSPG--SESTFEDFYCRVGLICVD -MRFVELA 162
Db 126 VAPEP...-PPTHVSPVTVGERSN...SVYUMFAPALALNWEHL-TUTKYA 176
QY 163 QIYVQKGCQLLVYFG-----AF-----NLTTGPAHWELLQSESAVGVVYVATASPA 229
Db 177 MYSMHEQVHVASWPGMSLYQPEVPAFGVDAQLIA-----IRMYALEGIVFVCTIQV 228
QY 210 -----RDKASYV--AMHSTVVPKGEVLAK- AGTEEAIVYGDIDKZLAEL 254
Db 229 VTPEHEFFCENEEOPLIGFGGFPAPILGPNRGRDLAIPLAUEEGILYADIDLSATLA 288
QY 255 PQOI-PVEPQKPSDLVAVEMKK 275
Db 289 QQAADPVGHYSRPPVLSINFNQ 310

RESULT 4
US-09-972-186A-2
; Sequence 2, Application US/09972186A
; Patent No. US20020137153A1
; GENERAL INFORMATION:
; APPLICANT: PAMER, SANDRA W.
; APPLICANT: HUISMAN, GJALT
; APPLICANT: MILLIS, JIM
; APPLICANT: SHELDON, ROGER
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: COX, ANTHONY
; APPLICANT: DAVIS, S. CHRISTOPHER
; TITLE OF INVENTION: ENANTIOSELECTIVE PRODUCTION OF AMINO CARBOXYLIC ACIDS
; FILE REFERENCE: 02-108110US
; CURRENT APPLICATION NUMBER: US/09/972,186A
; PRIOR FILING DATE: 2001-10-04
; PRICE APPLICATION NUMBER: 60/238,563
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide sequence
US-09-972-186A-2

Query Match 13.8%; Score 198.5; DB 10; Length 380;
Best Local Similarity 24.5%; Pred No. 3.2e-12;
Matches 79; Conservative 52; Mismatches 124; Indels 67; Gaps 15;

Fri Jan 31 13:48:38 2003

Patent No. 5824522
GENERAL INFORMATION:
APPLICANT: Ikenaka, Yasuhiro
APPLICANT: Nanba, Hirokazu
APPLICANT: Takano, Masayuki
APPLICANT: Yajima, Kazuyoshi
APPLICANT: Yamada, Yukio
APPLICANT: Takahashi, Satomi
APPLICANT: Okubo, Kazuma
APPLICANT: Yamada, Kazuhiko
APPLICANT: Hiraiishi, Yoshiko
TITLE OF INVENTION: Immobilized Enzyme Preparation and
TITLE OF INVENTION: Process for Producing D-a-Amino Acid
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/294,871A
APPLICATION NUMBER: US 07/917,111
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/971,758
FILING DATE: 12-APR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/917,111
FILING DATE: 07-AUG-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/211,641
FILING DATE: 11-APR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 400848/1990
FILING DATE: 07-DEC-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/01696
FILING DATE: 06-DEC-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 407922/1990
FILING DATE: 27-DEC-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 078840/1991
FILING DATE: 11-APR-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 140051/1991
FILING DATE: 12-JUN-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP92/00739
FILING DATE: 10-JUN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 212502/1992
FILING DATE: 10-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/01101
FILING DATE: 05-AUG-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 340078/1992
FILING DATE: 21-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/PCT NUMBER: 741274/121/AUPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

TELEX: 904136
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-294-871A-46
Query Match 16.8%; Score 241; DB 2; Length 303;
Best Local Similarity 27.1%; Pred. No. 2,26-19; Indels 40; Gaps 12;
Matches 75; Conservative 56; Mismatches 106;
QY 21 VTRACSFIREAATGGAKIVSLPECFNSPYGAK-YFFEYAE-----KIPGKSTQYISEV 72
Db 24 VVKLLDMLTKAASKGANFIVFPLALTLIFFRWYFTLAELESEFYETEMCPVVRPLPEY 83
QY 73 AKECST-YLIGGS--IFEDAGKLYNCAVFGDGTLLAYKRIHLFTIVPCGVITQ-- 127
Db 84 AAEGLGIGENLGYAEVLVVEGGVKKRPNTISLAVDESCHIVGFFYTHL---PRHFFYFAY 198
QY 128 -----ESKTLSPSD-SFSTFTTYGCVGSGIGCYDMFAELALIVACRATQILVYVGAFT 180
Db 139 RPFQHLKRYFPEGDLGFPVYVDAAKMMFELNURKWPPEAMVWGLRGAELICGGUN 196
QY 181 LTT-GPA-----HWELLQSRVDNQQVIVATASPAEDDKASTVANGSHSTVNPWG 229
Db 197 IPTHNESVEQHRLTSPHLLSWAGSYQTSAMSAAGKAGMFFHMYLCHGCTIVATG 265
QY 230 EVLAKAGT-EEAIVYSDIDLKLAELPQLIVFPPQKP 265
Db 456 EIVALTITLEDVITAADVLDRCNELEKEHIFHFQHP 292

RESULT 3
US-08-876-398A-46
Sequence 46, Application US/08876398A
Patent No. 6083752
GENERAL INFORMATION:
APPLICANT: IKENAKA, Yasuhiro
APPLICANT: NANBA, Hirokazu
APPLICANT: TAKANO, Masayuki
APPLICANT: YAJIMA, Kazuyoshi
APPLICANT: YAMADA, Yukio
APPLICANT: TAKAHASHI, Satomi
TITLE OF INVENTION: DNA CODING FOR DETAPRAMEYLASE IMPROVED IN
TITLE OF INVENTION: THERMOSTABILITY AND USE THEREOF
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,398A
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/211,641
FILING DATE: 11-APR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP93/01101
FILING DATE: 05-AUG-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 340078/1992
FILING DATE: 21-DEC 1992

QY	868	TGTTGCACAGAAATAGGACGATATGATTTCTTCAACATATCAACTCCCTATTAAATCTTT	927
Db	841	TOGTTTCACAGAAATAGGACGATATGATTTCTTCAACATATCAACTCCCTATTAAATCTTT	900
QY	928	AATGAAGAAAAAAA 942	
Db	901	TATGAAAAA 915	
RESULT 2			
LOCUS	BI333425		
DEFINITION	602996847F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE 513251 5', mRNA sequence.		
ACCESSION	BI333425		
VERSION	BI333425.1 GI:15018082		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates, Catarrhini; Homnidae; Homo.		
AUTHORS	1. (bases 1 to 893)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Sequencing by: Incyte Genomics, Inc. DNA Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11342 low: m column. 12 High quality sequence stop: 874.		
FEATURES			
source	1..893		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:513251"		
	/clone_lib="NIH_MGC_12"		
	/tissue_type="cervical carcinoma cell line"		
	/lab_host="DH10B"		
	/note="Organ: cervix, Vector: pCMV SP6P16; Site:1. NotI; Site:2. SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 Kb. Library prepared by Life Technologies."		
BASE COUNT	245 a 202 c 212 g 233 t		
ORIGIN			
Query Match	61.0%; Score 828.4; DB 13; Length 893;		
Best Local Similarity	97.6%; Pred. No. 2.7e-137;		
Matches	872; Conservative 0; Mismatches 17; Indels 4; Gaps 3		
QY	22	ATGACCTCTTCCGCTTGGGCTTCATCCAGCTTCAGATTTCTTCCATCAATCAGATAAC	81
Db	1	ATGACCTCTTCCGCTTGGGCTTCATCCAGCTTCAGATTTCTTCCATCAATCAGATAAC	60
QY	82	GTCCTCGCGCTTGTAGCTTCATCCGAGGAGGAGCAACGAGGAGCAAAATAGTTTCT	141
Db	61	GTCCTCGCGCTTGTAGCTTCATCCGAGGAGGAGCAACGAGGAGCAAAATAGTTTCT	119
QY	142	TTGCGGAAATGCTTTAAATCTCCATATGAGGAGAAATATTTCTGTAATGAGAGAAA	201
Db	120	TTGCGGAAATGCTTTAAATCTCCATATGAGGAGAAATATTTCTGTAATGAGAGAAA	179
QY	202	ATTCTGTGTAAATCTACACAGAGCTTCTGAGTAGCAAGAAATCCACATATATCTC	261
Db	180	ATTCTGTGTAAATCTACACAGAGCTTCTGAGTAGCAAGAAATCCACATATATCTC	239
QY	262	ATTGAGGCTCTATCCCTGGAAGAGATGCTGGGAAATATATATACACTGTGTGTTT	321

1

100

SOURCE human.
OPERATION
REFERENCE NIH-MGC http://mgi.nhl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bhs.fda.gov
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I M A G E Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I M A G E Consortium/LLNL at:
http://image.llnl.gov
plate: LLMI0174 row: a column: 21
High quality sequence stop: 710.
Location/Qualifiers
1. 919
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:4426652"
/clone_lib="NIH MGC 89"
/tissue_type="hypertrophoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
BASE COUNT 255 a 211 c 219 g 233 t 1 others
ORIGIN
Query Match 51.1%; Score 695; DB 12; Length 919;
Best Local Similarity 98.5%; Pred.No. 1.1e-113;
Matches 744; Conservative 0; Mismatches 5; Indels 6; Gaps 4;
QY 30 TTTCCGCTTGTATCCAGCTTCAGATTTCTCCATCAATCAGATAAGCTCAGTCG 89
Db 1 TTTCCGCTTGTATCCAGCTTCAGATTTCTCCATCAATCAGATAAGCTCAGTCG 60
QY 90 CGCTTGTAGCTTTCATCCGAGGAGGAGCAAGCAAGGAGGAGGAGGAGGAGGAG 149
Db 51 CGCTTGTAGCTTTCATCCGAGGAGGAGCAAGCAAGGAGGAGGAGGAGGAGGAG 120
QY 150 ATGCTTTTAACTTCATATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 209
Db 121 ATGCTTTTAACTTCATATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 210 TGAATCCACACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 269
Db 191 TGAATCCACACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 270 CTCTATCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 329
Db 241 CTCTATCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 299
QY 330 TGGAACTTTATAGCAAAAGATATAGAAAGATCCATCTGTTGACATTTGATGTTCCCTGGAAA 389
Db 300 TGGAACTTTATAGCAAAAGATATAGAAAGATCCATCTGTTGACATTTGATGTTCCCTGGAAA 359
QY 370 AATTAACATTTAAGAAATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTT 419
Db 360 AATTAACATTTAAGAAATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTT 419
QY 450 TCCCTTACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 509
Db 420 TCCCTTACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479
QY 510 AATCTACGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 569

Db 480 AATCTACGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
QY 570 TGAACACAGCCCATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 629
Db 540 TGAACACAGCCCATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
QY 630 GGGCCACAGGCTCTCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 689
Db 600 GGGCCACAGGCTCTCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
QY 690 CGTGTGTAACCCCTT-GGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 745
Db 660 CGTGTGTAACCCCTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719
QY 746 TGTATTTCAGACATAGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 720 TGTATTTCAGACATAGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 753

Search completed: January 31, 2003, 11:21:20
Job time : 2229 secs

581	QY	ATTGGAGTTACTTCAGCGGAAGCCGGCTGTGTGATAATCAGGTGTATGTGTGGCACACGCT	640
785	Db	ACTGGAGAGTGTTGCTGGGGCCGCTGCTATGCGAAGACAGTGTATATATATATATATATAT	844
641	QY	CTTCCTGCCCGGATACACAAAGCCTCCTATGTTGCTTGGGTATAGTATATATATATATAT	699
845	Db	CACAGTGTGGAGCCCACTATGAGAAAGAGAGAGATATATATATATATATATATATATAT	904
701	QY	CTTGGGGGGAGGTTCTAGCCAAAGTGGCACAGAAAGCAATGCTTATTATATATATATAT	760
905	Db	CGTGGGGAACATGGTGGCCCGGCTGTTGGAGAGAGAGAGCTGTGCTATATATATATAT	964
761	QY	ACCTGAAGAGTGTCCTGAATACCGACAGCAATCCCGCTTTTATACAGAAAGATATAT	820
965	Db	ACCTCAACTATCTGGACACATGGCCGGAACAGTGGCTATATATATATATATATATATAT	1024
821	QY	ACCTCTATG	829
1025	Db	ACCTCTATG	1033

RESULT 3

```

US-09-774-639-92
; Sequence 92, Application US/09774639
; Publication No. US20030003558A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: PZ013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: FA-1118 APPLICATION NUMBER: 09/214,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1165)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-774-639-92

```

Query Match 7.1%; Score 97; DB 9; Length 1203;
Best Local Similarity 50.9%; Pred No 9.7e-14;
Matches 256; Conservative 0; Mismatches 245; Indels 2; Gaps

QY	329	ATGGAACCTTTACTAGCAAAAGTATAGAAAGATCAATCTGTTTGACATTGATGTTCTCGGAA	388
Db	46	AAGGGGAGTAGTGGGCACTTACAGAGAGACACATCTGTGTGAGGATATATATTGATATG	164
QY	389	AAATTACATTTCAGAATCTCTA--AAACATTGATCCGGGTGATAGTTTCTCCACATTGGA	446
Db	156	AGGGGCTATGTGTGAAAGCAACTCTACCAATGCTGGAGCCAGTCTTGATGACATCTCAG	215
QY	447	TACTTCCTTACTGTCAGAGTGGGTCTGGGCATCTGCTACGACATGCGGTTTGGACAGATTGAC	506
Db	216	CACACCCAGCAGGCAAGATTTGTCCTAGCTGTCTGCTATGACATGAGTTTCTCCCTGAACTCTC	275
QY	507	ACAAATCTATGACACAGAGAGGCTGCCAGCTGTGGTATATCCAGAGAGCTTTTAAATCTGAC	566
Db	276	TCTGGCATTTGCTCAACTCTGAGCAGAGATACTTAGCTATCTCTCAGCTTTTGGATCCAT	335
QY	567	CACGTGACCAAGGCATATGGGAGATATTTACAGCAAGGCGGGCTGTGTATATCAATCAAGTATA	626
Db	336	TACAGGCGCCAGCCCACTGCGGAGGTGTGCTCGGGGCGGTGCTATCAAAATCAAGTGCTA	395
QY	627	TGTGGCCACAGGCTCTCTCTCCCGCGGATGTAACAAAGCTCTCTATGTTGCTCTGGGACACAG	686
Db	396	TGTAGTGTGACACAGACAGATGTGGACGCCCATATGACAAAGATCAAGTTTATGGCACAG	455

[illegible]

RESULTS :

```

1  RESULT 2
2  US-09 925-300-692
3  ? Sequence 692, Application US/09925300
4  ? Patent No US2002016161A1
5  ? GENERAL INFORMATION:
6  ? APPLICANT: Craig Rosen.
7  ? APPLICANT: Steve Ruben.
8  ? TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
9  ? FILE REFERENCE: PA01
10 ? CURRENT APPLICATION NUMBER: US/09/925,300
11 ? CURRENT FILING DATE: 2001-08-10
12 ? PRIOR APPLICATION NUMBER: PCT/US00/05988
13 ? PRIOR FILING DATE: 2000-03-08
14 ? PRIOR APPLICATION NUMBER: 60/124,270
15 ? PRIOR FILING DATE: 1999-03-12
16 ? NUMBER OF SEQ ID NOS: 1990
17 ? SOFTWARE: Patent In Ver. 2.0
18 ? SEQ ID NO 692
19 ? LENGTH: 1382
20 ? TYPE: DNA
21 ? ORGANISM: Homo sapiens
22 ? US-09-925-300-692

```

Query Match	7.3%	Score 99.8;	DB 10;	Length 1382;
Best Local Similarity	50.1%;	Prod No. 2.2e-14;		
Matches 275;	Conservative	0;	Mismatches 272;	Indels 2;
			Gaps	

[illegible]

Fri Jan 31 13:48:45 2003

```

Best Local Similarity 2.9%; Prod. No. 0.049;
Matches 11, Conservative 214; Mismatches 160; Indels 0; Gaps 0;

QY 702 GGGAGCTTCTTATGCTAAAGTGGCAGACAGAGCAATCGTGTATTCAGACATAGACCTG 765
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 765 AAGAACTGCTGAATACGCCAGCAATCCCGTTTATAGACAGAGCGATGACACCTC 825
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 826 TATGCTGTCAGATGAGAAAGCCCTAAAGTTATGTTCTTAATGTGTCACAGATAGGAC 885
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 886 CATATGATTCACACATAATCAATCCCTATTAAATTCITTAATGAAGAAAGAAATTT 945
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 946 AAAAAAAGAAAAACCTAGGCTTCTTATTCAGATGAGAAAGCCCTATTATGCTGACA 1005
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 1006 CTTTCCAGAGCAATTAAATAGTTAAAGAGATGACGCTGGAGCCAGAGAGAGAAAGC 1065
DL 1134 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
QY 1065 TGGCTGCTGTTTAAAGCTTTCTTCA 1070
DL 1074 EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

```

```

RESULT 2
US-09-415-946-2
; Sequence 2, Application US/09415946
; Patent No. 6376751
; GENERAL INFORMATION:
; APPLICANT: Sung, Z. Renee
; APPLICANT: Aubert, Dominique
; APPLICANT: Chen, Lingjing
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Reproductive Development in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 018941-00011005
; CURRENT APPLICATION NUMBER: US/09/415,946
; PRIOR FILING DATE: 1999-10-08
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 8648
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(3201)
; NAME/KEY: CDS
; LOCATION: join(4241..4335, 4448..4623, 4704..4823, 4903, 4956,
; OTHER INFORMATION: EMF1 gene and its promoter region
; NAME/KEY: exon
; LOCATION: (3202)..(3245)
; NAME/KEY: intron
; LOCATION: (3246)..(4159)
; NAME/KEY: exon
; LOCATION: (4160)..(4335)
; OTHER INFORMATION: exon 2
; NAME/KEY: intron
; LOCATION: (4336)..(4447)
; NAME/KEY: exon
; LOCATION: (4448)..(4623)

```

```

; OTHER INFORMATION: exon 3
; NAME/KEY: intron
; LOCATION: (4624)..(4703)
; NAME/KEY: exon
; LOCATION: (4704)..(4823)
; OTHER INFORMATION: exon 4
; NAME/KEY: intron
; LOCATION: (4824)..(4902)
; NAME/KEY: exon
; LOCATION: (4903)..(4956)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (4957)..(5045)
; NAME/KEY: exon
; LOCATION: (5046)..(6307)
; OTHER INFORMATION: exon 6
; NAME/KEY: intron
; LOCATION: (6308)..(6447)
; NAME/KEY: exon
; LOCATION: (6448)..(8065)
; OTHER INFORMATION: exon 7
; NAME/KEY: intron
; LOCATION: (8066)..(8300)
; NAME/KEY: exon
; LOCATION: (8301)..(8648)
; OTHER INFORMATION: exon 8
; US-09-415-946-2

Query Match 3.11; Score 42.8; LG 4; Length 8648;
Best Local Similarity 54.48; Prod. No. 0.077;
Matches 86; Conservative 72; Indels 0; Gaps 0;

QY 829 GCTGTGAGATGAAAAAGCCCTAAAGTTTAAATGTTCTTAATGTCACAGATAGCAGAT 888
DL 1748 GCTGTGAGATGAAAAAGCCCTAAAGTTTAAATGTTCTTAATGTCACAGATAGCAGAT 1807
QY 889 ATGATTTCTACACATATCACTCCCTATTAAATTTCTTAATGAGAAAAAATTTAAA 948
DL 1808 TAATTCGAAAGTAGCTCAACATCCCAATTAATTCCTTTATCAATATCGAATATGCAA 1867
QY 949 AAAAAAAGAAAAACCTAGGTTCTTCTATTGAGATGAGA 986
DL 1868 AGGCTTAAGAAAAAAGTTTCTTTTGTGTAA 1905

RESULT 3
US-09-415-946-1
; Sequence 1, Application US/09415946
; Patent No. 6376751
; GENERAL INFORMATION:
; APPLICANT: Sung, Z. Renee
; APPLICANT: Aubert, Dominique
; APPLICANT: Chen, Lingjing
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Reproductive Development in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 018941-00011005
; CURRENT APPLICATION NUMBER: US/09/415,946
; PRIOR FILING DATE: 1999-10-08
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 17341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: genomic DNA (Ecotype Columbia) from CD92 clone
; NAME/KEY: promoter
; LOCATION: (1)..(3201)
; NAME/KEY: CDS
; LOCATION: join(4241..4335, 4448..4623, 4704..4823, 4903, 4956,

```

GenCore version 5.1.3
Copyright (c) 1998 - 2003 GenCore Ltd.

OM protein - protein search, using sw model

Run 8 January 31, 2003, 13:48:43, Search time 4 seconds
(without alignments)
1672 618 Million cell updates/sec

Title: us_09_855_294b_1

Perfect score: 1438

Sequence: 1 VTSEFLATQVQVGNVSTN... QVSEFQVQVAVEMPT 276

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 67540 seqs, 1684715 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DR seq length: 6

Maximum DR seq length: 276

Post-processing: Minimum Match 60

Maximum Match 100%

Listing first 48 summaries

Database: SPTRMBL.21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_plage.*
- 10: sp_plant.*
- 11: sp_ricet.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviro.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1438	100.0	276	Q9NQR4	Geng4 homo sapien
2	1434	99.7	276	Q8WU60	Q8wuf6 homo sapien
3	1322	91.9	276	Q9JHW2	Q9jhw2 mus musculus
4	1074	74.7	231	Q9DQ93	Q9dq93 mus musculus
5	914	65.0	180	Q9CTG9	Q9ctg9 mus musculus
6	822.5	57.2	369	Q9CJF8	Q9cjf8 arabidopsis
7	814	56.6	318	Q9LYH1	Q9lyh1 arabidopsis
8	781.5	54.3	293	Q9VEH4	Q9veh4 iridopis
9	712	49.5	327	Q97839	Q97839 methanobact
10	485.5	33.8	327	Q76091	Q76091 homo sapien
11	483	33.6	290	Q9CBF2	Q9cbf2 mus musculus
12	483	33.6	323	Q9S526	Q9s526 mus musculus
13	483	33.6	323	Q9PIN4	Q9pin4 mus musculus
14	479	33.2	288	Q91930	Q91930 ven f-8 145
15	475	33.0	323	Q9VDK1	Q9vdk1 mus musculus
16	475	33.0	590	Q9WJ79	Q9wj79 myogin 145

Fri Jan 31 13:48:43 2003

```

QY 51 IPGSESTOKLSEVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
DL 51 IPGSESTOKLSEVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
QY 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
DB 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
QY 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
DB 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
QY 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276
DB 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276

RESULT 2
AC Q9JHW2 PRELIMINARY; PRT; 276 AA.
DE 01-MAR-2002 (TrEMBLrel. 20, created)
DI 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 21, last annotation update)
DE Nit protein 2.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP STRAUSBERG P.;
RS SUBMITTED (JAN-2002) to the EMBL/GenBank/DBJ databases.
RE FMPT, P000795, AAR00153;
DE Intespro; IPR003010, Ntlse/Chydtse.
DR Pfam; PF00795; CN hydrolase; 1.
SQ SEQUENCE 276 AA; 36580 MW; 10AA/97601E7BR4C CPC64;

Query March 99.78; Score 1434; DB 4; Length 276;
Best Local Similarity 99.68; Pred. No. 3.8e-124;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTSFLALIQICISIKSDNVTRACSFIREATOCAGKIVSLPECFNSPYGAKYFPEVAEK 60
DB 1 MTSFLALIQICISIKSDNVTRACSFIREATOCAGKIVSLPECFNSPYGAKYFPEVAEK 60
QY 61 IPRSTQKLSVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
DB 61 IPRSTQKLSVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
QY 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
DB 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
QY 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
DB 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
QY 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276
DB 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276

RESULT 3
AC Q9JHW2 PRELIMINARY; PRT; 276 AA.
DE 01-MAR-2002 (TrEMBLrel. 15, created)
DI 01-MAR-2002 (TrEMBLrel. 15, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Nit protein 2 (1190017B10pik protein) (PIKEN cDNA 1190017B10

```

```

DE gene).
GN NIT 09 1190017B10pik
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=1095938;
RA Pace H.C., Hodawadekar S.C., Draganescu A., Huang J., Biedanowski P.,
RA Pekarsky Y., Croce C.M., Brenner C.;
RT "Crystal structure of the worm NitFhit Rosetta Stone protein reveals a
RT Nit tetramer binding two Fhit dimers.";
RL Curr. Biol. 10:907-917(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa Y., Iizawa M., Nishi Y., Fiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Resole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Roffelli D., Bojunga N., Carinci P., de Ronald M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons F., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Shioh K. F.,
RA Suzuki H., Toyooka K., Wang Y.H., Weir G., Whittaker C., Wilming B.,
RA Wyshaw-Bolis A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL, AC284573; AAR0102 1;
DR ENBL, AC004535; BAB23354 1;
DR ENBL, BC020553; AAR00153 1;
DR MGD; MGI:1913477; 1190017B19Rik.
DR Intespro; IPR003010, Ntlse/Chydtse.
DR Pfam; PF00795; CN hydrolase; 1.
SQ SEQUENCE 276 AA; 36580 MW; 74CFCM449793D6D CPC64;

Query March 91.58; Score 1322; DB 11; Length 276;
Best Local Similarity 89.58; Pred. No. 8.3e-114;
Matches 247; Conservative 17; Mismatches 12; Indels 0; Gaps 0;

QY 1 MTSFLALIQICISIKSDNVTRACSFIREATOCAGKIVSLPECFNSPYGAKYFPEVAEK 60
DB 1 MTSFLALIQICISIKSDNVTRACSFIREATOCAGKIVSLPECFNSPYGAKYFPEVAEK 60
QY 61 IPRSTQKLSVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
DB 61 IPRSTQKLSVAKECSYVLIGSSIFEDAGKLNITCAVEFGDGTLLAKYKIHLPFDIV 120
QY 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
DB 121 PKKITFOESKTLSPGDSFSTFTTTCYRVLGICGYDMRFELAQIYAQPCQLLVPGAFN 180
QY 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
DB 181 LTGTGAHWELLQSRSPAVDQNVVATASPARDDKASVAVGHSSTVNVFWGEVLAKAGTEEA 240
QY 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276
DB 241 IVYSIDILKZLAFIPQCIQVFPQPPSTLYAVEMKPP 276

```


Fri Jan 31 13:48:43 2003

```

DR PROGITH, PS01227; UPF0012; 1
SQ SEQUENCE 327 AA, 3596 MW, 99788904BA62781 CFC64;

Query Match
33.6%; Score 483; DB 11; Length 290;
Best Local Similarity 38.3%; Pred. No. 2.4e-36;
Matches 107; Conservative 54; Mismatches 106; Indels 14; Gaps 6;

QY 2 TSFELALQLQIQLSSSI--KSNVITPACNSFIEAATQAKTVSLSECEISFYCAVYFPE - 56
DB 5 TSWELPLVAVCVSTSTPNKQENFKTCABLVQEAARLGACLAFLPEAFD--FIARNPAETL 62
QY 57 YAKIPIQESTPLSEVAFESIVLIGRSIFE--EFAKUNTCVAFVPGDGLLAFY 110
DB 63 LLSFELNDLLQVSSQLAFENSIM--SDFHEPDSGWEKGFVIRWHLIRFSEVAVY 122
QY 111 EKIHLEDLIVKRTIFQESKTLSTGDSFS--TPTTPYVFWLIGCTTMDPFAELAIQVABG 169
DB 123 KATHLCVGLVIFSGGVRRESNITFGTILEFTVPTACVGLAICVDMFPFELSILAIAG 182
QY 170 COLLVTPGAKRLITTPALWELLORSFAVTCVTVATASAFETVACVAVWGHSTVWVNG 229
DB 183 ABILTYPSAFESVTVFAHWEVILLRKAIESQCTIVAAAGCGRHRETRACYTHGVVYVW 242
QY 230 EYLAKEASTEEA--VYGDIDLPKLAEIPQCTIPVFFPFSSILY 269
DB 243 TVVRESSEFSLCLAFILULHFLGCMFQHLFVFGHREFTLY 282

RESULT 12
O88526 PRELIMINARY; PRT; 323 AA.
AC O88526;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Nitrilase homolog 1.
GN NIT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RF MEDLINE-98337986; PubMed-9671749;
RX Pekarsky Y., Campiglio M., Siprashvili Z., Druck T., Sedkov Y.,
RA Tillib S., Draganesu A., Wermuth P., Rothman J.H., Huebner K.,
KA Bachberg A.M., Mazo A., Brenner C., Croce C.M.;
RT "Nitrilase and fhit homologs are encoded as fusion proteins in
KT drosophila melanogaster and caenorhabditis elegans.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8744-8749(1998).
DR EMBL; AF069985; AAC40184.1; -
DR MGD; MGI:1350916; Nitr1.
DR InterPro; IPR003010; Nitrse/CNhydrtse
DR InterPro; IPR001110; UPF0012.
DR Pfam; PF00795; CN hydrolase; 1.
DR PROSITE; PS01227; UPF0012; 1
SQ SEQUENCE 323 AA, 35717 MW, 10151CEBL51DF2C7 CFC64;

Query Match
33.6%; Score 483; DB 11; Length 323;
Best Local Similarity 38.3%; Pred. No. 2.4e-36;
Matches 107; Conservative 54; Mismatches 105; Indels 14; Gaps 6;

QY 2 TSFELALQLQIQLSSSI--KSNVITPACNSFIEAATQAKTVSLSECEISFYCAVYFPE - 56
DB 38 TSWELPLVAVCVSTSTPNKQENFKTCABLVQEAARLGACLAFLPEAFD--FIARNPAETL 95
QY 57 YAKIPIQESTPLSEVAFESIVLIGRSIFE--EFAKUNTCVAFVPGDGLLAFY 110
DB 63 LLSFELNDLLQVSSQLAFENSIM--SDFHEPDSGWEKGFVIRWHLIRFSEVAVY 122
QY 111 EKIHLEDLIVKRTIFQESKTLSTGDSFS--TPTTPYVFWLIGCTTMDPFAELAIQVABG 169
DB 123 KATHLCVGLVIFSGGVRRESNITFGTILEFTVPTACVGLAICVDMFPFELSILAIAG 215
QY 170 COLLVTPGAKRLITTPALWELLORSFAVTCVTVATASAFETVACVAVWGHSTVWVNG 229
DB 183 ABILTYPSAFESVTVFAHWEVILLRKAIESQCTIVAAAGCGRHRETRACYTHGVVYVW 242

us-09-855-294b-1.1.rsp
33.6%; Score 485.5; DB 4; Length 327;
Best Local Similarity 36.7%; Pred. No. 1.7e-36;
Matches 105; Conservative 46; Mismatches 98; Indels 37; Gaps 6;

QY 6 LALIQQLQSSIFSTIVTACSPFPAAT--TAKIVSEPECF-----NSPYGAK 52
DB 49 VAVCVSTSTPNKQENFKTCABLVQEAARLGACLAFLPEAFD--FIARNPAETLHSELOK 108
QY 57 YAKIPIQESTPLSEVAFESIVLIGRSIFE--EFAKUNTCVAFVPGDGLLAFY 107
DB 63 LLSFELNDLLQVSSQLAFENSIM--SDFHEPDSGWEKGFVIRWHLIRFSEVAVY 122
QY 111 EKIHLEDLIVKRTIFQESKTLSTGDSFS--TPTTPYVFWLIGCTTMDPFAELAIQVABG 169
DB 123 KATHLCVGLVIFSGGVRRESNITFGTILEFTVPTACVGLAICVDMFPFELSILAIAG 215
QY 170 COLLVTPGAKRLITTPALWELLORSFAVTCVTVATASAFETVACVAVWGHSTVWVNG 229
DB 183 ABILTYPSAFESVTVFAHWEVILLRKAIESQCTIVAAAGCGRHRETRACYTHGVVYVW 242

QY 224 VWPWGVENAKATEAIVSDIDLPKLAEIPQCTIPVFFPFSSILY 269
DB 274 VWPWGVENAKATEAIVSDIDLPKLAEIPQCTIPVFFPFSSILY 319

RESULT 11
Q9DBF2 PRELIMINARY; PRT; 290 AA.
AC Q9DBF2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Nitrilase 1.
GN NIT1.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RF MEDLINE-98337986; PubMed-9671749;
RX Kawai J., Shinagawa A., Shibata Y., Konno H., Adachi J., Fukuda S.,
RA Arakawa T., Hara A., Fukunishi Y., Kiyokawa H., Kondo S., Yamanaoka I.,
KA Aizawa K., Izawa M., Nishi K., Kiyokawa H., Kasukawa T., Saito F.,
RT "Nitrilase and fhit homologs are encoded as fusion proteins in
KT drosophila melanogaster and caenorhabditis elegans.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8744-8749(1998).
DR EMBL; AF069985; AAC40184.1; -
DR MGD; MGI:1350916; Nitr1.
DR InterPro; IPR003010; Nitrse/CNhydrtse
DR InterPro; IPR001110; UPF0012.
DR Pfam; PF00795; CN hydrolase; 1.
DR PROSITE; PS01227; UPF0012; 1
SQ SEQUENCE 290 AA, 31896 MW, 34493DBF52170C71 CFC64;

```


us-09-855-294b-1.rspt

Fri Jan 31 13:48:43 2003

```

QY 57 -YAKIPGESTKLGFAVARECCSIYLLIGGIFE. ...EDAGKLYMTCAVEFGPGTLLAKY 110
DL 96 LUSEFINGDLGJYSZLAKECJIKWLEJOFHEFGQWECQKTYNCHVLNSKGSWASY 155
QY 111 RKIHLFDINDVPGKITPOESKTI.SPDSES-TEOTPYCRVGLGICYDMRFAELAQIYAQEG 169
DL 156 RETHLCTVPIPGQWFESENUTPFGCTI.FFVVTFAKVGCLATCYDMRPFELS.KLAQAG 215
QY 179 QGLLVYFGAFNLTTGPAHWELLGCRSAVTNQVYVATASPARCDKASYVAMGHSITVVNFWG 229
DB 216 ABILTYSSAFSGSVTGPAAWELILFAPATIESQCQVIAAQQGRHHETASVGHSMVVDPMG 275
QY 230 EVLAKAGTEEAIVVSDIDIKFLAEIPQOIPVFPQKSDLY 269
DB 276 TVVARQSEGPGLAPITLHFLQCMFCHI.PVFQHRRPDLY 315

```

Search completed: January 31, 2003, 07:52:17
 b time : 38 secs

3094	3121	449	<800	1236	1188	1186	1188
1051	1038	5421	5341	2940	2953	2608	2566
1997	2009	7500	7792	462	<800	67	<800
1757	1712	9225	9161	2948	2953	872	945
2095	2140	12	<800	1377	1335	3925	3879
3662	3697	256	<800	6	<800	2592	2566
225	<800	871	803	3543	3569	14998	14913
5089	5000	470	<800	1182	1107		
1274	1264	1367	1370	3453	3433		
2598	2666	574	<800	15967	16220		
3984	3998	10682	10819	574	<800		
2929	2967	9515	8485	3627	3569		
748	<800	3904	2992	120	<800		
1911	1868	2600	2656	7795	7733		
1985	2009	1196	1181	2469	2566		
3807	3697	459	<800	1351	1335		
1410	1391	459	<800	1857	1844		
1536	1532	2189	2219	5216	5137		
8613	8718	3769	3753	8339	8206		
9224	9191	221	<800	7068	6952		
731	<800	7852	7792	3968	3879		
159	<800	6996	6955	1155	1133		
733	<800	11277	10819	232	<800		
1003	1038	2022	2015	4622	4519		
11724	13758	177	<800	244	<800		
6749	6707	4807	4794	5628	5560		
906	896	328	<800	2755	2774		
1342	1264	2023	2219	11617	11404		
123	<800	3702	3753	1552	1188		
2167	2140	2960	2992	5663	5560		
438	<800	30494	32561	2922	2953		
5264	5226	1011	1046	931	945		
2069	2080	7193	2219	442	<900		
3535	3697	2741	2749	478	<800		
739	<800	195	<800	1123	1071		
4300	4342			857	945		
<hr/>							
FEATURES							
source		Location/Qualifiers					
		1..142816					
		/organism="Homo sapiens"					
		/db xref="taxon:9606"					
		/chromosome="1"					
		/clone="RP11-156K6"					
		/clone_lib="RPC1 human BAC library 11"					
<hr/>							
Query Match 45 18; Score 748 4; DB 9; Length 142816;							
Best Local Similarity 78 88; Pref No. 3 7e-176;							
Matches 1050; Conservative 0; Mismatches 326; Indels 56; Gaps 119							
QY	2	TGTTGTTTGTTCAGATTCATGACCTCTTTTGGTTTGGGTTTATAGTTTCAATTTT	61				
DB	35580	TGTTGCTTCCAG	35521				
QY	62	TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	121				
DB	35520	TTT-CATCAAAATCAGATAAACAATGAGATTTGATCTTCTGAGATTTGATCTTCTG	35462				
QY	122	AAGGAGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	191				
DB	35461	AAGGAGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	35402				
QY	192	TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	241				
DB	35401	TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	35342				
QY	242	AGGAATGAGAT	301				
DB	35341	AGGAAT	35288				
QY	302	ATAACATCTGT	361				
DB	35287	ATAACATCTGT	35228				
QY	362	ATCTTTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	421				
DB	35227	ATCTTTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	35169				
QY	422	AGGAGAT	481				
DB	35167	AGGAGAT	35109				
QY	482	AGGAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	541				

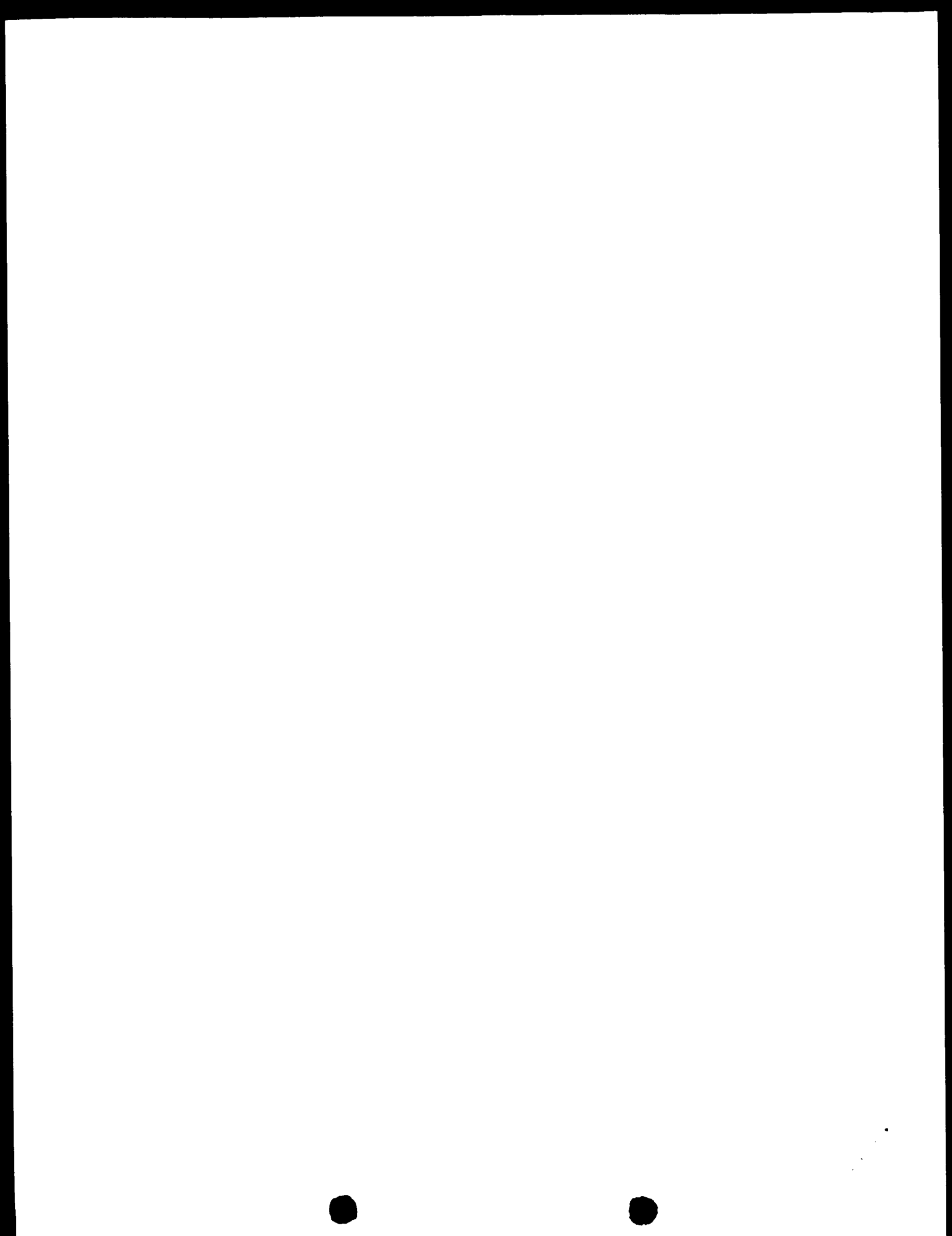
Db 865 AAGAAAGCTTTGATATCTGTTCAAAAATGTCACCAACAGGATGATGCTCTG 915

KEYWORDS

AX330407.1 GI:18103383

VERSION

KEYWORDS



XX The present invention relates to human hNit3-ase (nitrilase) protein,
CC which is expressed in Cushings adrenal gland tumour, and its coding
CC sequence (AA04214 and AA09084). The present invention also relates to a
CC preparation method of said protein and nucleic acid sequence, and a
CC method of detecting human hNit3-ase nucleic acid sequence and polypeptide
CC in sample.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1438; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.5e-149;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSFRLALIQIQTSSIKSDNVTPACSFIREAATOGAKIVSLPECFNSPYGAKYFPEYAEK 60
DB 1 MTSFRLALIQIQTSSIKSDNVTPACSFIREAATOGAKIVSLPECFNSPYGAKYFPEYAEK 60

QY 61 IPGESTOKLSEVAKECSIYLGSIPEEDAGKLYNTCAVFGDGTLLAKYRKIHLPDIDV 120
DB 61 IPGESTOKLSEVAKECSIYLGSIPEEDAGKLYNTCAVFGDGTLLAKYRKIHLPDIDV 120

QY 121 PGKITFOESKTLSPGDSFSTFDTPCVGHGICVDMPPFAELAQIYAORGCOLLYVPGAEN 180
DB 121 PGKITFOESKTLSPGDSFSTFDTPCVGHGICVDMPPFAELAQIYAORGCOLLYVPGAEN 180

QY 181 LTTGPAHWELLQSRVAVDNGVYVATASPARDDKASYVAVWGHSTVVPWGEVLAKAGTEEA 240
DB 181 LTTGPAHWELLQSRVAVDNGVYVATASPARDDKASYVAVWGHSTVVPWGEVLAKAGTEEA 240

QY 241 IVTSDIDLKLAE:EQQIPVFPQPPSELYAVEMFKP 276
DB 241 IVTSDIDLKLAE:EQQIPVFPQPPSELYAVEMFKP 276

RESULT 2
AA076477
ID AA076477 standard; Protein; 276 AA.
AC AA076477;
XX
XX 23-SEP-2002 (first entry)
XX Novel human nitrilase-like protein.
XX
XX Nitrilase, enzyme, human, NHP; gene therapy.
XX
XX Homo sapiens.
XX
XX W0200194566-A2.
XX
XX 13-DEC-2001
XX
XX 25 JAN 2001; 2001WO-US-6687
XX
XX 28 JAN 2000; 2000US-178000P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Gencho G, Hilban E, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX
XX WFI; 2002-519061/55.
XX N-PSDB; ABN83908.
XX
XX Novel polynucleotides encoding novel human proteins sharing sequence
PT similarity with nitrilase proteins, useful for generating primers, and
PT probes which are used in identification of targets for drug discovery
PT
XX
XX (Claim 4; Page 29-30; 36pp; English.
PS
XX The present sequence is the protein sequence of a novel human

CC protein (NHP) that shares structural similarity with nitrilase
CC enzymes from a wide variety of organisms. The invention provides
CC nitrilase-like NHPs (see AB076477-83) and polynucleotides (see
CC ABN83908-15), host cell expression systems, fusion proteins,
CC antibodies, antagonists, agonists, and other compounds that
CC modulate NHP expression or activity. Those can be used to
CC detect mutant or inappropriately expressed NHPs for the diagnosis
CC of disease, or to screen for drugs effective in the treatment of
CC perturbations of the normal function of an NHP. Engineered host
CC cells and/or transgenic animals producing NHP can be used to
CC identify compounds that bind the endogenous NHP receptor and to
CC identify compounds that trigger NHP-mediated activities or
CC pathways.
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1438; DB 23; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.5e-149;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSFRLALIQIQTSSIKSDNVTPACSFIREAATOGAKIVSLPECFNSPYGAKYFPEYAEK 60
DB 1 MTSFRLALIQIQTSSIKSDNVTPACSFIREAATOGAKIVSLPECFNSPYGAKYFPEYAEK 60

QY 61 IPGESTOKLSEVAKECSIYLGSIPEEDAGKLYNTCAVFGDGTLLAKYRKIHLPDIDV 120
DB 61 IPGESTOKLSEVAKECSIYLGSIPEEDAGKLYNTCAVFGDGTLLAKYRKIHLPDIDV 120

QY 121 PGKITFOESKTLSPGDSFSTFDTPCVGHGICVDMPPFAELAQIYAORGCOLLYVPGAEN 180
DB 121 PGKITFOESKTLSPGDSFSTFDTPCVGHGICVDMPPFAELAQIYAORGCOLLYVPGAEN 180

QY 181 LTTGPAHWELLQSRVAVDNGVYVATASPARDDKASYVAVWGHSTVVPWGEVLAKAGTEEA 240
DB 181 LTTGPAHWELLQSRVAVDNGVYVATASPARDDKASYVAVWGHSTVVPWGEVLAKAGTEEA 240

QY 241 IVTSDIDLKLAE:EQQIPVFPQPPSELYAVEMFKP 276
DB 241 IVTSDIDLKLAE:EQQIPVFPQPPSELYAVEMFKP 276

RESULT 3
AA015790
ID AA015790 standard; Protein; 276 AA.
AC AA015790;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human Nit2 protein.
XX
XX Human; Nit2; cytostatic; neuroprotective, cellular pathway; therapy;
KW apoptosis; proliferative disorder; degenerative disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 2..12 /label= NS1_strand
FT Region 16..34 /label= NH1_helix
FT Region 37..42 /label= NS2_strand
FT Region 54..74 /label= NH2_helix
FT Region 78..89 /label= NS3_strand
FT Region 90..101 /label= NS4_strand
FT Region 106..111 /label= NS5a_strand
FT Region 116..119 /label= NS5b_strand
FT

Region 123..127
 /label= NS6a_strand
 Region 140..143
 /label= NS6b_strand
 Region 146..151
 /label= NS7_strand
 Region 153..169
 /label= NH3_helix
 Region 171..178
 /label= NS8_strand
 Region 182..199
 /label= NH4_helix
 Region 201..213
 /label= NS9_strand
 Region 215..226
 /label= NS10_strand
 WO200187959-A2.
 22-NOV-2001.
 15-MAY-2001; 2001WO-US15664.
 16-MAY-2000; 2000US 204713P.
 (UJJE-) UNIV JEFFERSON THOMAS.
 Croce C, Brenner C, Pekarski Y;
 WPI; 2002-092994/11.
 N PSDB; AAD24256.

Isolated cDNA encoding human, mouse, frog and yeast Nit2 proteins, useful to find molecules that mimic or antagonize Phit interaction for the treatment of proliferative or degenerative diseases.

Claim 5; Page 53-54; 51pp; English.

The invention relates to isolated nucleic acids comprising a fully defined cDNA nucleotide sequence encoding human, Xenopus laevis and mouse Nit2 proteins. Nit and Phit proteins are encoded as fusion proteins in invertebrates and as separate polypeptides in vertebrates. Nit and Phit interact physically and functions in same cellular pathways. Molecules which bind Nit2 and mimic or antagonise Phit interaction are used to treat diseases in which activity of Nit2 protein is altered in a mammal. Phit mimics induce apoptosis and are particularly useful to treat proliferative disorders, whilst Phit antagonists promote cell proliferation and are particularly useful to treat degenerative disease. The present sequence is human Nit2 protein.

XX Sequence 276 AA;

Query Match 100.0%; Score 1438; DB 23; Length 276.
 Best Local Similarity 100.0%; Pred No 1 5e-149;
 Matches 276; Conservative 3; Mismatches 0; Indels 0; Gaps 0.

QY 1 MTSPFLALIGLOSSISSENVVTAACSIPIPAATAGAKIVSLPEFNSFYGAKYFEVAEK 60
 IL 1 MTSPFLALIGLOSSISSENVVTAACSIPIPAATAGAKIVSLPEFNSFYGAKYFEVAEK 60
 QY 61 IPSESTQPLSEVAPERSIYLIGDSIHAFEDAFVNTAVENFQETLLAKYKHLPEIDV 120
 EL 61 IPSESTQPLSEVAPERSIYLIGDSIHAFEDAFVNTAVENFQETLLAKYKHLPEIDV 120
 QY 121 PPKITPQESTLSPGSPSTPDTPVQPVLSGIVYMPFAELAQIYAPGQGLLYVPSAFN 180
 DB 121 PPKITPQESTLSPGSPSTPDTPVQPVLSGIVYMPFAELAQIYAPGQGLLYVPSAFN 180
 QY 181 LTTGPAHMLLQSPAVNQQVYVATASPADDKASYVAVGSHSTVYVNWGEVLAKAGTEA 240
 DB 181 LTTGPAHMLLQSPAVNQQVYVATASPADDKASYVAVGSHSTVYVNWGEVLAKAGTEA 240
 QY 241 IVYSQIDLLFLAEIPGQIPVFRGFESLYAVENKXP 276

DB 241 IVYSQIDLLFLAEIPGQIPVFRGFESLYAVENKXP 276

RESULT 4

ABB76483
 ID ABB76483 standard; Protein; 285 AA.

XX AC ABB76483;

XX 23 SEP-2002 (first entry)

XX Novel human nitrilase-like protein.

XX Nitrilase; enzyme; human; NHP; gene therapy

XX Homo sapiens.

XX WO200194566 A2.

XX 13-DEC-2001.

XX 25-JAN-2001; 2001WO US266687.

XX 28-JAN-2000; 2000US 179000P.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Hilbun E, Scoville J, Turner CA, Friedrich G, Abuin A;
 XX Zambrowsicz E, Sands AT;

XX WPI; 2002-519061/55.
 XX N-PSDB; ABH83814.

XX Novel polynucleotides encoding novel human proteins sharing sequence similarity with nitrilase proteins, useful for generating primers, and probes which are used in identification of targets for drug discovery

XX Disclosure; Page 34-35; 36pp; English.

The present sequence is the protein sequence of a novel human protein (NHP) that shares structural similarity with nitrilase enzymes from a wide variety of organisms. The invention provides nitrilase-like NHPs (see ABB76477-83) and polynucleotides (see ABB76480-16), host cell expression systems, fusion proteins, antibodies, antagonists, agonists, and other compounds that modulate NHP expression or activity. These can be used to detect mutant or inappropriately expressed NHPs for the diagnosis of disease, or to screen for drugs effective in the treatment of perturbations of the normal function of an NHP. Engineered host cells and/or transgenic animals producing NHP can be used to identify compounds that bind the endogenous NHP receptor and to identify compounds that trigger NHP-mediated activities or pathways.

XX Sequence 285 AA;

Query Match 99.4%; Score 1247; DB 23; Length 285;
 Best Local Similarity 97.1%; Pred No 1 4e-143;
 Matches 285; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MTSPFLALIGLOSSISSENVVTAACSIPIPAATAGAKIVSLPEFNSFYGAKYFEVAEK 60
 DB 1 MTSPFLALIGLOSSISSENVVTAACSIPIPAATAGAKIVSLPEFNSFYGAKYFEVAEK 60
 QY 61 IPSESTQPLSEVAPERSIYLIGDSIHAFEDAFVNTAVENFQETLLAKYKHLPEIDV 120
 DB 61 IPSESTQPLSEVAPERSIYLIGDSIHAFEDAFVNTAVENFQETLLAKYKHLPEIDV 120
 QY 121 PPKITPQESTLSPGSPSTPDTPVQPVLSGIVYMPFAELAQIYAPGQGLLYVPSAFN 180
 DB 121 PPKITPQESTLSPGSPSTPDTPVQPVLSGIVYMPFAELAQIYAPGQGLLYVPSAFN 180
 QY 181 LTTGPAHMLLQSPAVNQQVYVATASPADDKASYVAVGSHSTVYVNWGEVLAKAGTEA 240
 DB 181 LTTGPAHMLLQSPAVNQQVYVATASPADDKASYVAVGSHSTVYVNWGEVLAKAGTEA 240

useful to find molecules that mimic or antagonize Phit interaction for the treatment of proliferative or degenerative diseases

Claim 11; Page 54; 61pp; English.

The invention relates to isolated nucleic acids comprising a fully defined cDNA nucleotide sequence encoding human, Xenopus laevis and mouse Nit2 proteins. Nit and Phit proteins are encoded as fusion proteins in invertebrates and as separate polypeptides in vertebrates. Nit and Phit interact physically and functionally in same cellular pathways. Molecules which bind Nit2 and mimic or antagonise Phit interaction are used to treat diseases in which activity of Nit2 protein is altered in a mammal. Phit mimics induce apoptosis and are particularly useful to treat proliferative disorders, whilst Phit antagonists promote cell proliferation and are particularly useful to treat degenerative disease. The present sequence is mouse Nit2 protein.

Sequence 276 AA;

Query March 91.9%; Score 1322, DB 23; Length 276;
Best Local Similarity 89.5%; Pred. No. 8.8e-137;
Matches 247; Conservative 17; Mismatches 12; Indels 0; Gaps 0;

QY 1 MTSFRLALIQLOISSIKSDNVTRACSFIREAATQGAIVSLPECNPSFYGAKYFPEVAEK 60
DB 1 MSTFRLALIQLOVSSIKSLENTTACSLVPEAAFGANIVSLPEFRSFYGTTFPEVAEK 60
QY 61 IPGESTOKLSEVAKECSIYLLIGSSIPEEDAGKLYNTCAVEGEGTLLAKYFKHLEFDIV 120
DB 61 IPGESTOKLSEVAKECSIYLLIGSSIPEEDAGKLYNTCAVEGEGTLLAKYFKHLEFDIV 120
QY 121 FGKITFQESKTLSPGDSFTFDTPYCVGLGICYDMRFAELAQIYAORGCOLLYVPGAFN 180
DB 121 FGKITFQESKTLSPGDSFTFDTPYCVGLGICYDMRFAELAQIYAORGCOLLYVPGAFN 180
QY 181 LTTGPAHWELLQSRANDVNVYVATASPARDDKASYVANGHSTVVFNGEVLAKAGTEEA 240
DB 181 LTTGPAHWELLQSRANDVNVYVATASPARDDKASYVANGHSTVVFNGEVLAKAGTEEA 240
QY 241 IVYSIDILKFLAEIPQIPVPPQPPSDLYAVEMK 276
DB 241 ILYSIDILKFLAEIPQIPVPPQPPSDLYAVEMK 276

RESULT 6

ABB63710
ID ABB63710 standard; Protein; 283 AA.

AC ABB63710;

XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17922.

XX Drosophila; developmental biology; cell signalling, insecticide;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2001; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

QY 181 LTTGPAHWELLQSRANDVNVYVATASPARDDKASYVANGHSTVVFNGEVLAKAGTEEA 240
DB 181 LTTGPAHWELLQSRANDVNVYVATASPARDDKASYVANGHSTVVFNGEVLAKAGTEEA 240
QY 241 IVYSIDILKFLAEIPQIPVPPQPPSDLYAVEMK 276
DB 241 ILYSIDILKFLAEIPQIPVPPQPPSDLYAVEMK 276

RESULT 5

AAE15791
ID AAE15791 standard; Protein; 276 AA.

XX AAE15791;

DT 26-MAR-2002 (first entry)

XX Mouse Nit2 protein

DE Mouse; Nit2; cytostatic; neuroprotective; cellular pathway; therapy;
XX apoptosis; proliferative disorder; degenerative disease.

OS Mus sp.

XX Key Location/Qualifiers

FT Region 2..12 /label= NSA1_strand

FT Region 16..34 /label= NH1_helix

FT Region 37..42 /label= NS2_strand

FT Region 54..74 /label= NH2_helix

FT Region 78..89 /label= NS3_strand

FT Region 90..101 /label= NS5a_strand

FT Region 116..119 /label= NS5b_strand

FT Region 123..127 /label= NS6a_strand

FT Region 140..143 /label= NS6b_strand

FT Region 146..152 /label= NS7_strand

FT Region 159..169 /label= NH3_helix

FT Region 171..178 /label= NS8_strand

FT Region 181..199 /label= NH4_helix

FT Region 201..213 /label= NS9_strand

FT Region 216..226 /label= NS10_strand

XX WC200187958-A2.

XX 22-NOV-2001

XX 15-MAY-2001; 2001WO-US15654.

XX 16-MAY-2000; 2000US-204713P.

XX (UWJE-) UNIV JEFFERSON THOMAS.

XX Croce C, Brenner C, Pekarski Y;

XX WPI; 2002-082984/11.

XX N-PSDB; AAD25457.

XX Isolated cDNA encoding human, mouse, frog and yeast Nit2 proteins,

PN WO200187958-A2.
 XX 22-NOV-2001.
 PD 25-MAY-2001; 2001WO-US15664.
 XX 16-MAY-2000; 2000US-204713P.
 PF 16-MAY-2000; 2000US-179000P.
 PR (UYJE-) UNIV JEFFERSON THOMAS.
 XX Croce C, Brenner C, Pekarski Y;
 XX WPI: 2002-082984/11.
 XX Isolated cDNA encoding human, mouse, frog and yeast Nit2 proteins,
 PT useful to find molecules that mimic or antagonize Phit interaction for
 PT the treatment of proliferative or degenerative diseases -
 XX Claim 17; Page 56-57; 61pp; English.
 XX The invention relates to isolated nucleic acids comprising a fully
 defined cDNA nucleotide sequence encoding human, Xenopus laevis and
 mouse Nit2 proteins. Nit and Phit proteins are encoded as fusion
 proteins in invertebrates and as separate polypeptides in vertebrates.
 CC Nit and Phit interact physically and functions in same cellular pathways.
 CC Molecules which bind Nit2 and mimic or antagonise Phit interaction are
 CC used to treat diseases in which activity of Nit2 protein is altered in
 CC a mammal. Phit mimics induce apoptosis and are particularly useful to
 CC treat proliferative disorders, whilst Phit antagonists promote cell
 CC proliferation and are particularly useful to treat degenerative disease.
 CC The present sequence is yeast Nit3 protein.
 XX Sequence 291 AA;
 SQ Query Match 48.3%; Score 694 5; DB 23, Length 291;
 Best Local Similarity 49.8%; Pred. No. 1.3e-67;
 Matches 137; Conservative 52; Mismatches 77; Indels 9; Gaps 4;
 QY 5 PLALIGLQIIS- IKSDNVTRACSFIPRAATQ--GAKIVSLPECFNSPYGAKYFPEYAE-- 59
 DB 17 FVALVQLGSGSPDQWANIQPAATFIERAMKQPOTKUVLPECFNSPYSTDFRKYSEVI 71
 QY 60 --KIPGESTQKLEVAKPCGTYLGGSTPEED--AGKLYNTCAVPGPGDTLLAKYRKHL 115
 DB 72 NPKSESTVSQVPLENLANKFKIILVASTTPELDPKTKIKYNTSIIFNEDGKLDKRRKVHL 131
 QY 116 FQIDVPKHTTQESKTHSPGDSFSTFDTGYFVGIGICVDWRFELAQIYAKRGQQLLVY 175
 DB 132 FQVDIPNGISPHSESTLSPGKESTTIDTKYKFRVSI CYDWRFPPELMLSAKKGAFAMLY 191
 176 PGAFNLTTGPAHWELLQSRVNDNGVYVATASPARDDKAGYVAVGHSITVVPFWGVLAKA 235
 DB 192 PSAFTVTGSLHLLHLLASPAVDNQQVYMLCSPPARNLQSSVHAYGHSITVDPGRKIVAE 251
 QY 236 STEEAIVSSDINPFPLAIEGQIIVFPCKRSLYA 270
 DB 232 QEGEEIIIAEELPEVIESPPQAVPIITKGPFPDVS 286
 RESULT 12
 ID ABB76480 standard; Protein; 168 AA.
 XX ABB76480;
 AC ABB76480;
 XX 23-SEP-2002 (first entry)
 DE Novel human nitrilase-like protein.
 XX Nitrilase; enzyme, human, NHP; gene therapy.
 XX Homo sapiens.
 XX WO200187958-A2.
 XX

PN WO200194566-A2.
 XX 13-DEC-2001.
 XX 25-JAN 2001; 2001WO-US26687.
 PF 28-JAN-2000; 2000US-179000P.
 PR (LEXI-) LEXICON GENETICS INC.
 XX Donoho G, Hilbun E, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 XX WPI: 2002-519061/55.
 DR N-PSDB; ABN83811.
 DR Novel polynucleotides encoding novel human proteins sharing sequence
 XX similarity with nitrilase proteins, useful for generating primers, and
 PT probes which are used in identification of targets for drug discovery
 PT Disclosure; Page 32; 36pp; English.
 XX The present sequence is the protein sequence of a novel human
 CC protein (NHP) that shares structural similarity with nitrilase
 CC enzymes from a wide variety of organisms. The invention provides
 CC nitrilase-like NHPs (see ABB76477-83) and polynucleotides (see
 CC ABN83808-15), host cell expression systems, fusion proteins, that
 CC antibodies, antagonists, agonists, and other compounds that
 CC modulate NHP expression or activity. These can be used to
 CC detect mutant or inappropriately expressed NHPs for the diagnosis
 CC of disease, or to screen for drugs effective in the treatment of
 CC perturbations of the normal function of an NHP. Engineered host
 CC cells and/or transgenic animals producing NHP can be used to
 CC identify compounds that bind the endogenous NHP receptor and to
 CC identify compounds that trigger NHP-mediated activities or
 CC pathways.
 XX Sequence 168 AA;
 SQ Query Match 44.3%; Score 646; DB 23, Length 168;
 Best Local Similarity 93.5%; Pred. No. 1.2e-62;
 Matches 123; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 142 ETFYVPVLGICVYMPFARLAIYVAGPQDGLLVTFEAFNLTTGPAHWELLQSRVNDNGV 201
 DB 25 ETAYCRYGLGICVDMRFARLAIYVAGPQDGLLVTFEAFNLTTGPAHWELLQSRVNDNGV 84
 QY 202 YVATASPARDDKASVAVGHSITVVPFWGVLAKAGTAEAVVSDIDLKLAEPQOIPVF 144
 DB 85 YVATASPARDDKASVAVGHSITVVPFWGVLAKAGTAEAVVSDIDLKLAEPQOIPVF 144
 QY 260 PQPSSDLYAVEMK 274
 DB 145 RQKNIFLNMQRK 157
 RESULT 13
 ID ABB76479 standard; Protein; 121 AA.
 XX ABB76479;
 AC ABB76479;
 XX 23-SEP-2002 (first entry)
 DE Novel human nitrilase-like protein.
 XX Nitrilase; enzyme, human, NHP; gene therapy.
 XX Homo sapiens.
 XX WO200194566-A2.
 XX

13-DEC-2001
 25-JAN-2001; 2001WO 0526687
 28 JAN-2000; 2000US 170000P
 (LEXI-) LEXICON GENETICS INC.
 Donoho G, Hilbun E, Scoville J, Turner CA, Friedrich G, Abuin A;
 Zambrowicz B, Sands AT;
 WPI: 2002-519061/55
 N-PSDB; ABR83810.
 Novel polynucleotides encoding novel human proteins sharing sequence
 similarity with nitrilase proteins, useful for generating primers, and
 probes which are used in identification of targets for drug discovery
 Disclosure; Page 31; 36pp; English.
 The present sequence is the protein sequence of a novel human
 protein (NHP) that shares structural similarity with nitrilase
 enzymes from a wide variety of organisms. The invention provides
 nitrilase-like NHPs (see AB76477-83) and polynucleotides (see
 ABR83808-15), host cell expression systems, fusion proteins,
 antibodies, antagonists, agonists, and other compounds that
 modulate NHP expression or activity. These can be used to
 detect mutant or inappropriately expressed NHPs for the diagnosis
 of disease, or to screen for drugs effective in the treatment of
 perturbations of the normal function of an NHP. Engineered host
 cells and/or transgenic animals producing NHP can be used to
 identify compounds that bind the endogenous NHP receptor and to
 identify compounds that trigger NHP-mediated activities or
 pathways.
 Query Match 43.8%, Score 627, DP 23, Length 121, Indels 0, Gaps 0;
 Best Local Similarity 100.0%; Pred. No. 9e-61;
 Matches 121, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 156 MFPAELAGTIVACQPSLLVTEGAPUTTTGAWWELLQPSFVEMQVAVATASAPDFAS 215
 1 MPFAELAGTIVACQPSLLVTEGAPUTTTGAWWELLQPSFVEMQVAVATASAPDFAS 60
 216 YVAMGHSVVNPNWGEVIAKAGTAEIVYSIDILKLLAEIRQQIPVFPQKSDLYAVEMKK 120
 61 YVAMGHSVVNPNWGEVIAKAGTAEIVYSIDILKLLAEIRQQIPVFPQKSDLYAVEMKK 120
 276 P 276
 121 P 121
 RESULT 14
 ABR76481
 ID ABR76481 standard; Protein; 130 AA
 AC ABR76481;
 DT 23-SEP-2002 (first entry)
 DE Novel human nitrilase-like protein.
 KW Nitrilase; enzyme; human; NHP; gene therapy.
 OS Homo sapiens.
 XX WO200104566-A2.
 13-DEC-2001

25-JAN-2001; 2001WO 0526687
 28 JAN-2000; 2000US 170000P
 (LEXI-) LEXICON GENETICS INC.
 Donoho G, Hilbun E, Scoville J, Turner CA, Friedrich G, Abuin A;
 Zambrowicz B, Sands AT;
 WPI: 2002-519061/55
 N-PSDB; ABR83812.
 Novel polynucleotides encoding novel human proteins sharing sequence
 similarity with nitrilase proteins, useful for generating primers, and
 probes which are used in identification of targets for drug discovery
 Disclosure; Page 33; 36pp; English.
 The present sequence is the protein sequence of a novel human
 protein (NHP) that shares structural similarity with nitrilase
 enzymes from a wide variety of organisms. The invention provides
 nitrilase-like NHPs (see AB76477-83) and polynucleotides (see
 ABR83808-15), host cell expression systems, fusion proteins,
 antibodies, antagonists, agonists, and other compounds that
 modulate NHP expression or activity. These can be used to
 detect mutant or inappropriately expressed NHPs for the diagnosis
 of disease, or to screen for drugs effective in the treatment of
 perturbations of the normal function of an NHP. Engineered host
 cells and/or transgenic animals producing NHP can be used to
 identify compounds that bind the endogenous NHP receptor and to
 identify compounds that trigger NHP-mediated activities or
 pathways.
 Query Match 39.8%, Score 573, DP 23, Length 130;
 Best Local Similarity 93.3%; Pred. No. 9e-55;
 Matches 117, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 156 MFPAELAGTIVACQPSLLVTEGAPUTTTGAWWELLQPSFVEMQVAVATASAPDFAS 215
 1 MPFAELAGTIVACQPSLLVTEGAPUTTTGAWWELLQPSFVEMQVAVATASAPDFAS 60
 216 YVAMGHSVVNPNWGEVIAKAGTAEIVYSIDILKLLAEIRQQIPVFPQKSDLYAVEMKK 274
 61 YVAMGHSVVNPNWGEVIAKAGTAEIVYSIDILKLLAEIRQQIPVFPQKSDLYAVEMKK 274
 RESULT 15
 AAY68739
 ID AAY68739 standard; Protein; 464 AA.
 AC AAY68739;
 DT 05-MAY-2000 (first entry)
 DE Amino acid sequence of the Nhl1 protein.
 KW Nhl1 gene; nitrilase; tumor suppressor gene; PHT; chromosome 3p14.2;
 FRAB; cancer; genome allele inactivation.
 OS Homo sapiens.
 OS Mus sp.
 OS Drosophila melanogaster.
 OS Caenorhabditis elegans.
 XX Key Location/Qualifiers
 PH Misc-difference 6
 FT /note- "unspecified amino acid encoded by TNT"
 XX WO200003685-A2.

Fri Jan 31 13:48:37 2003

```

PD 27-JAN-2000.
XX
XX
PF 20-JUL-1999; 99W0-NS16366.
XX
XX
PR 20-JUL-1998; 98US-0093350.
XX
XX
PA (OYJE-) UNIV JEFFERSON THOMAS.
XX
XX
PI Croce CM;
XX
XX
WP1: 2000-171195/15.
DR N-PSDB; AAZ46101.
XX
XX
PT Novel nitrilase homologs used as diagnostic and therapeutic reagents
PT for the detection and treatment of cancer -
XX
XX
PS Disclosure; Fig 6; 25pp; English.
XX
XX
CC The present sequence is encoded by the coding region of human, murine,
CC Drosophila melanogaster and Caenorhabditis elegans NIT1 gene. The
CC human and mouse NIT1 genes are members of an uncharacterised
CC mammalian gene family with homology to bacterial and plant nitrilases.
CC The tumour suppressor gene FHIT in D. melanogaster and C. elegans code
CC for fusion proteins in which the FHIT domain is fused with a Nit domain.
CC In mouse and humans, FHIT and NIT are encoded by two different genes,
CC localised on chromosomes 3 and 1 in human and 14 and 1 in mouse. The
CC human FHIT gene at chromosome 3p14.2, spanning the constitutive
CC chromosomal fragile site FRA3B, is often altered in most common forms
CC of human cancer. The NIT1 protein overcomes the mutated inactivation
CC of the genome alleles. The NIT1 genes, encoded polypeptides, derivatives
CC and analogues of them, and antibodies are used as diagnostic and
CC therapeutic reagents for the detection and treatment of cancers.
XX
XX
SQ Sequence 464 AA;

```

Query Match 33.8%; Score 485.5; DB 21; Length 464;
Best Local Similarity 36.7%; Pred. No. 2.9e-44;
Matches 105; Conservative 46; Mismatches 98; Indels 37; Gaps 6;

```

QY 6 LALICLQISSKSDNVTACRFIREAATGGAKIVSLPECF--.....NSPYGAK 52
Db 84 VAVCVVTSIPURQNFETCAEIVPAAPIGACIAPFAFTFIARDPAETLHLSEPLGK 143
QY 53 YFPEYAEKIPGESTOKLSEVAKECSIYLIIGSGIPE-----EDAGKLYNTCAVFGPDGILL 107
Db 144 LLEEV-----TOLARECCGLWLSLGGFHERGQDWEQTKIYNCHVLLNSKGAV 191
QY 138 AYYPYTHLEPIIVPGKIFQESKTLSPGDSF-STFDTPTVCFVGI.GICYDMRPAELAQIYA 166
Db 132 ATYKATHLCDVEIPQGGMCESNSTMPGDSLESVPSTPAGKIGLAVCYDMRPFELSALA 251
QY 167 QPQGGHMYVYAFNITTPAHWELQRPAPVONQVYVATASPA---RDDKASYVAWGHT 223
Db 252 QAGAILLAYPSAFESITTPAHWEVLIAPAIFETQYVVAQAACGPHHKKRASY---GHSM 308
QY 224 VVNFQGEVLAKATTEAIVYSIDLEKLAIEIQQIPVPQKPSDLY 269
Db 309 VVDFWGTVAQCSEGPGLTAPIDLNLYIPQRPPLHLPVFPQHPRPDLY 354

```

Search completed: January 31, 2003, 07:51:17
Job time : 39 secs

XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX Claim 1; SEQ ID NO 869; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in AB343533-AB397455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at www.int/pub/published/pct_sequences.

Query Match 34.7% Score 472.5 SE 24 Length 499
 Best Local Similarity 99.2% Pred No. 1 le-110;
 Matches 485; Conservative 0; Mismatches 5; indels 1; GAPS 1

[illegible]

RESULT 9
ABL62579/C
ID ABL62579 standard; DNA; 489 RP
XX
AC ABL62579;
XX

[illegible]

RESULT 3	
ABN94371/C	
ID ABN94371	standard; DNA; 489 BP.
XX	
XX	
AC	ABN94371;
XX	
XX	13-AUG-2002 (first entry)
DE	Gene #869 used to diagnose liver cancer.
XX	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW	metastatic liver tumour; cytostatic; expression profile; disease state;
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.

XX	OS	Homo sapiens.
XX	XX	
PN	WQ000224103-A2	
XX	XX	
PD	11-APR-2002.	
XX	XX	
PF	02-OCT-2001; 2001WO-018305R9.	
XX	XX	
PR	02-OCT-2000; 2000US-237954P.	
XX	XX	
PA	(GENE-) GENE LOGIC INC.	
XX	XX	
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX	XX	
DR	WPI; 2002-426119/45.	

RESULT 14

ABN83812
ID ABN83812 standard, cDNA, 393 bp.

AC ABN83812:

03-SEP-2005

XX
polynucleotide encoding novel human nitrilase-like protein

XX MA class of human gene therapy gene ss

xv Homomorphisms

XX	Key	Location/Qualifiers
PH	CDS	1..393
FT		/*tag- a /product= "NHP"

PN W0200194566-A3

13-DEC-2001

XX DE JAN 2001 1801NOV86C097

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX Doncho G, Hilbun E, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B. Sands AT:
PI

WPI: 2002-519061/55.

DR P-PSDB; ABB76481.

Novel polynucleotides encoding novel human proteins sharing sequence similarity with nitrilase proteins, useful for generating primers, and probes which are used in identification of targets for drug discovery

PS Disclosure: page 32: 36pp: English

The present sequence is the coding sequence of an isolated nucleic acid molecule encoding a novel human protein (NHP) that shares structural similarity with nitrilase enzymes from a wide variety of organisms. The invention provides nitrilase-like NHPs (see ABB76477-83) and polynucleotides encoding them (see ABB93808-15), host cell expression systems, fusion proteins, antibodies, antagonists, agonists, and other compounds that modulate NHP expression or activity and are useful for diagnosis, drug screening, clinical trial monitoring and the treatment of physiological disorders. The NHP polynucleotide sequences were compiled from gene trapped cDNAs and clones isolated from human prostate, lymph node, pituitary, mammary gland and kidney cDNA libraries. NHP polynucleotide sequences can be used in the treatment of diseases involving NHP, for detecting mutant NHPs or inappropriately expressed NHPs for the diagnosis of disease, and for screening drugs effective in treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. Nucleic acid fragments can be used as primers in PCR to screen libraries, isolate clones and prepare cloning and sequencing templates. They can also be used as hybridisation probes for screening libraries and for assessing gene expression patterns, as well as for identification, selection and validation of novel molecular targets for drug discovery. Labeled NHP probes can be used to screen a genomic library to identify polymorphisms, determine the genomic structure of a given locus/allele and design diagnostic tests. The probes also have use in defining and monitoring drug action and toxicity. NHP oligonucleotides can be used to design antisense, ribozyme and/or triple helix molecules. Addressable arrays comprising the NHP polynucleotides can be used to identify and

